Elementary Inclusion Relations for Generalized Numerical Ranges

Dedicated to Olga Taussky Todd

Moshe Goldberg and E. G. Straus†
Department of Mathematics
University of California, Los Angeles, California

ABSTRACT

Let $\gamma_1,\ldots,\gamma_n$ be complex constants. The set $W_{(\gamma_1,\ldots,\gamma_n)}(A)=\{\Sigma\gamma_j(Ax_j,x_j)\}$, where $(x_1,\ldots,x_n)$ vary over all orthonormal systems in $\mathbf{C}^n$, is called a generalized numerical range of a given $n\times n$ matrix A. In this paper we study inclusion relations of the form $W_{(\gamma_1,\ldots,\gamma_n)}\subset\lambda W_{(\gamma_1',\ldots,\gamma_n')}$ which hold uniformly for all n-square matrices A. In particular we concentrate on the case where the coefficients are real. Such inclusion relations yield simple inequalities among generalized numerical radii. Finally, a further generalization of the above numerical range is discussed.

1. INTRODUCTION

Let A be an $n\times n$ complex matrix, let $c=(\gamma_1,\ldots,\gamma_n)\in\mathbf{C}^n$ be a fixed complex vector, and let Λ_n be the set of all orthonormal n-tuples of vectors in $\mathbf{C}^n$. In this paper we study some inclusion relations between generalized numerical ranges which are sets in the complex plane of the form

$$W_c(A)=W_{(\gamma_1,\ldots,\gamma_n)}(A)=\left\{\sum_{j=1}^{n}\gamma_j(Ax_j,x_j):(x_1,\ldots,x_n)\in\Lambda_n\right\}.$$

From the definition it is clear that $W_c(A)$ actually depends only on the *unordered* set $\{\gamma_1,\ldots,\gamma_n\}$ rather than on the ordered n-tuple $c=(\gamma_1,\ldots,\gamma_n)$. In the following the vector c will always stand as a representative of the set $\{\gamma_1,\ldots,\gamma_n\}$, and we write $c\sim c'$ if c and c' represent the same set.

We recall now the definition of the k-numerical range given by Halmos [1, Sec. 167], which after a simple normalization becomes

$$W_k(A)=\left\{\frac{1}{k}\operatorname{tr}(PAP):P=\text{projection of rank }k\right\}\qquad(1\leqslant k\leqslant n).$$

†The research of the first author was sponsored in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-76-3046. The work of the second author was supported in part by NSF Grant MPS 71-2884.

Evidently $W_k(A)$ may be written as

$$W_k(A)=\left\{\frac{1}{k}\sum_{j=1}^{k}(Ax_j,x_j):(x_1,\dots,x_k)\in\Lambda_k\right\}, \tag{1.1a}$$

where Λ_k is the set of all k-tuples of orthonormal vectors in $\mathbf{C}^n$. Hence we see that

$$W_k(A)=W_c(A)\qquad\text{with}\quad c=\frac{1}{k}(e_1+\dots+e_k), \tag{1.1b}$$

$\{e_j\}_{j=1}^n$ being the standard basis for $\mathbf{C}^n$. Thus, the k-numerical range is a special case of a generalized numerical range. In particular, for $k=1$, i.e., for $c=e_1$, we obtain the classical range

$$W(A)=W_1(A)=\{(Ax,x):|x|=1\}.$$

It is also clear that

$$W_n(A)=\left\{\frac{1}{n}\operatorname{tr}A\right\}.$$

Berger [1, Sec. 167] has shown that $W_k(A)$ is convex. It was later proven by Westwick, [2], that $W_c(A)$ is convex for any $c\in\mathbf{R}^n$. Westwick also gave an example which shows that for complex vectors $c\in\mathbf{C}^n$ with $n\geqslant 3$, the range $W_c(A)$ may fail to be convex.

Certain inclusion relations involving k-numerical ranges were given in [3]. As in [3], we are interested here in inclusions which hold uniformly for all $A\in\mathbf{C}_{n\times n}$, that is, for all $n\times n$ complex matrices. In this paper we shall restrict our attention to elementary inclusion relations, i.e., relation of the simple form

$$W_c(A)\subset\lambda W_{c'}(A),\qquad \lambda=\text{constant}. \tag{1.2}$$

In a forthcoming paper we shall consider inclusion relations involving finite linear combinations and integrals of generalized numerical ranges.

We begin in Sec. 2 with some definitions. This leads, in Sec. 3, to the construction of inclusion relations of type (1.2) for the general case $c,c'\in\mathbf{C}^n$. Further results are obtained in Sec. 4 for the case $c,c'\in\mathbf{R}^n$. In Sec. 5, we derive some inequalities among generalized numerical radii. Finally, in Sec. 6, we define a further, and in a certain sense an ultimate, generalization of the concept of numerical range.

2. PARTIAL ORDER RELATIONS

We begin by defining two partial order relations among complex vectors.

DEFINITION 1. (i) For $c=(\gamma_1,\dots,\gamma_n)$ and $c'=(\gamma_1',\dots,\gamma_n')$ in $\mathbf{C}^n$, we say that $c \prec c'$ if there exists a doubly stochastic matrix S (i.e., a matrix with non-negative entries whose row and column sums are 1), such that $c=Sc'$.

(ii) The vector c is obtained from c' by *pinching* if two components γ_i', γ_j' of c' are replaced by γ_i, γ_j with

$$\gamma_i=\alpha\gamma_i'+(1-\alpha)\gamma_j', \quad \gamma_j=(1-\alpha)\gamma_i'+\alpha\gamma_j', \qquad 0\leqslant\alpha\leqslant 1, \tag{2.1}$$

while the other components of c remain unchanged. Note that pinching an n-tuple c' consists of moving two of its components towards their midpoint, and thus decreasing

$$\operatorname{conv}(c')\equiv \text{convex hull}\{\gamma_1',\dots,\gamma_n'\}.$$

A similar concept of pinching was used in [4] by Horn and Steinberg.

(iii) We say that $c \prec\prec c'$ if c is obtained from c' by a succession of a finite number of pinchings.

Note that the relations $\prec$, $\prec\prec$ are in fact relations between the *unordered* n-tuples $\{\gamma_1,\dots,\gamma_n\}$ and $\{\gamma_1',\dots,\gamma_n'\}$. In case (i) it follows from the fact that doubly stochastic matrices are closed under multiplications by permutation matrices. For case (iii) it follows directly from the definition.

THEOREM 1. *The relation $c\prec\prec c'$ implies $c\prec c'$, but not conversely.*

Proof. If $c\prec\prec c'$, then assume for simplicity that c has been obtained from c' by a single pinch. Hence, for some $i,j\in\{1,\dots,n\}$ and α with $0\leqslant\alpha\leqslant 1$, we have (2.1). So $c=Sc'$, where S is the doubly stochastic matrix defined by

$$S_{pq}=\begin{cases} 1, & p=q\neq i,j,\\ \alpha, & (p,q)=(i,i),(j,j),\\ 1-\alpha, & (p,q)=(i,j),(j,i),\\ 0 & \text{otherwise.}\end{cases}$$

Consequently $c\prec c'$, and the first part of the proof is established.

Next consider the vectors $c=(\frac{1}{2},i/2,\frac{1}{2}+i/2)$ and $c'=(0,1,i)$. Clearly

$$c=Sc' \qquad \text{with} \quad S=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \tag{2.2}$$

so $c \prec c'$. However the components of c are all located on the different edges of $\operatorname{conv}(c)$. Therefore, any chain of non-trivial pinches on c' yields a vector c'', where at least two components of c are outside $\operatorname{conv}(c'')$. Hence $c \neq c''$ and the relation $c \prec\prec c'$ fails to hold. ■

We now wish to show that $\prec$ is a partial order relation. For this purpose we need the next lemma, which seems of independent interest.

LEMMA 1. *If $c \prec c'$ and $c' \prec c$, then $c \sim c'$.*

Proof. Let $\alpha_1,\dots,\alpha_k$ be the distinct components of c, ordered so that $|\alpha_1| \geqslant \cdots \geqslant |\alpha_k|$. Let the multiplicity of α_l be m_l ($\sum m_l = n$), and assume that c has been arranged to take the form

$$c=(\alpha_1,\dots,\alpha_1,\dots,\alpha_k,\dots,\alpha_k). \tag{2.3}$$

In view of the remark following Definition 1, the relations $c \prec c'$, $c' \prec c$ are still valid; hence there exist doubly stochastic matrices S, S′ such that

$$c=Sc' \quad \text{and} \quad c'=S'c; \tag{2.4}$$

thus $c=SS'c$. Since the class of $n\times n$ doubly stochastic matrices form a multiplicative semigroup, we have

$$c=Tc \qquad (T=SS'), \tag{2.5}$$

where T is doubly stochastic as well. We assert that

$$T=T_1 \oplus \cdots \oplus T_k, \tag{2.6}$$

where T_l is doubly stochastic of order $m_l \times m_l$.

To prove (2.6) assume for simplicity that $k=2$, i.e.,

$$c=(\alpha_1,\dots,\alpha_1,\alpha_2,\dots,\alpha_2), \qquad \alpha_1\neq\alpha_2, \quad |\alpha_1|\geqslant|\alpha_2|,$$

where the multiplicity of α_l $(l=1,2)$ is m_l, and $m_1+m_2=n$. Take any of the first m_1 components of the equality in (2.5), say the ith one. Since $|\alpha_1|\geqslant|\alpha_2|$, this leads to

$$|\alpha_1|=\left|\left(\sum_{j=1}^{m_1}T_{ij}\right)\alpha_1+\left(\sum_{j=m_1+1}^{n}T_{ij}\right)\alpha_2\right|$$

$$\leqslant\left(\sum_{j=1}^{m_1}T_{ij}\right)|\alpha_1|+\left(\sum_{j=m_1+1}^{n}T_{ij}\right)|\alpha_2|\leqslant\left(\sum_{j=1}^{n}T_{ij}\right)|\alpha_1|=|\alpha_1|,$$

Hence we have equality, which, in view of the fact that $\alpha_1\neq\alpha_2$, may hold if and only if $T_{ij}=0$ for $j=m_1+1,\dots,n$. This means that the first m_1 rows of T vanish beyond their m_1 entry, so all the weight of these rows is concentrated in the first m_1 columns. Consequently, the first m_1 columns of T vanish beyond their m_1 element as well, and we obtain the desired decomposition $T=T_1\oplus T_2$.

Next recall that doubly stochastic matrices are convex combinations of permutation matrices P_σ. In particular, $S=\sum_\sigma\alpha_\sigma P_\sigma$; thus

$$T=SS'=\sum_\sigma\alpha_\sigma P_\sigma S'.$$

The matrices $\alpha_\sigma P_\sigma S'$ in the above sum have non-negative entries; hence they must all have the same block decomposition as T. Now we choose a coefficient α_τ with $\alpha_\tau\neq0$, and conclude that $P_\tau S'$ decomposes according to (2.6). Since $P_\tau S'$ is doubly stochastic and it has the same decomposition (2.3) as c, it follows that $P_\tau S'c=c$. So, finally, by (2.4),

$$c'=S'c=\left(P_\tau^{-1}\right)\left(P_\tau S'c\right)=P_\tau^{-1}c\sim c,$$

and the lemma follows. ∎

Remark . The above proof contains a special case of the following observation on group-rings over the reals (or any ordered field). Let

$$R(G)=\{\Sigma\alpha_i g_i:\alpha_i\in\mathbf{R},\ g_i\in cg\}$$

be a group-ring of G over $\mathbf{R}$, and let K_G be the convex hull of G in $R(G)$, that is,

$$K_G=\{\Sigma\alpha_i g_i:\alpha_i\geqslant0,\ \Sigma\alpha_i=1\}.$$

Then K_G is a multiplicative semigroup whose units are the elements of G. If H is a subgroup of G, then K_H is a sub-semigroup of K_G, and two elements u, v of K_G satisfy $uv \in K_H$ if and only if there exists an element $g \in G$ such that ug and $g^{-1}v$ are in K_H. Thus the only divisors, in K_G, of elements of K_H are associates of elements of K_H.

We conclude this section with the following property of $\prec$ and $\prec\prec$.

THEOREM 2. *The relations* $\prec$ *and* $\prec\prec$ *are partial order relations on the set of unordered n-tuples.*

Proof. We have to show that $\prec$ and $\prec\prec$ are reflexive, transitive and antisymmetric. The first two properties are easily verified, and by Theorem 1, $c \prec\prec c'$ implies $c \prec c'$. So it suffices to prove the antisymmetry of $\prec$, i.e., that $c \prec c'$ together with $c' \prec c$ yields $c \sim c'$. But this is the statement of Lemma 1, and the proof is complete. ■

3. ELEMENTARY INCLUSION RELATIONS

Before considering a general $n \times n$ case, we present the following result concerning 2×2 matrices.

LEMMA 2. *If A is a 2×2 matrix, then for any α_1, α_2,*

$$W_{(\alpha_1,\alpha_2)}(A) = (\alpha_1 - \alpha_2) W\left(A - \tfrac{1}{2}(\operatorname{tr} A) I\right) + \tfrac{1}{2}(\alpha_1 + \alpha_2)\{\operatorname{tr} A\}. \tag{3.1}$$

Thus $W_{(\alpha_1,\alpha_2)}(A)$ is convex.

Proof. As before, let Λ_2 denote the set of all orthonormal pairs of 2-vectors. If x_1, x_2 is in Λ_2, then

$$\begin{aligned}
&\alpha_1 (Ax_1, x_1) + \alpha_2 (Ax_2, x_2) \\
&= \tfrac{1}{2}(\alpha_1 - \alpha_2)(Ax_1, x_1) - \tfrac{1}{2}(\alpha_1 - \alpha_2)(Ax_2, x_2) + \tfrac{1}{2}(\alpha_1 + \alpha_2)\left[(Ax_1, x_1) + (Ax_2, x_2)\right] \\
&= (\alpha_1 - \alpha_2)(Ax_1, x_1) - \tfrac{1}{2}(\alpha_1 - \alpha_2)\operatorname{tr} A + \tfrac{1}{2}(\alpha_1 + \alpha_2)\operatorname{tr} A \\
&= (\alpha_1 - \alpha_2)\left(\left[A - \tfrac{1}{2}(\operatorname{tr} A) I\right] x_1, x_1\right) + \tfrac{1}{2}(\alpha_1 + \alpha_2)\operatorname{tr} A.
\end{aligned} \tag{3.2}$$

So, (3.1) is obtained from (3.2) as x_1, x_2 vary over Λ_2.

The convexity of $W_{(\alpha_1,\alpha_2)}(A)$ is implied by the convexity of the (classical) numerical range and the lemma follows. ■

Using the above lemma we obtain our first general inclusion relation.

COROLLARY 1. *If (γ_1,γ_2) is obtained from (γ_1',γ_2') by pinching, then*

$$W_{(\gamma_1,\gamma_2)}(A)\subset W_{(\gamma_1',\gamma_2')}(A) \qquad \forall A\in \mathbf{C}_{2\times 2}. \tag{3.3}$$

Proof. By definition of pinching there exists an α, $0\leqslant\alpha\leqslant 1$, such that

$$\gamma_1=\alpha\gamma_1'+(1-\alpha)\gamma_2', \qquad \gamma_2=(1-\alpha)\gamma_1'+\alpha\gamma_2'.$$

Hence, by Lemma 2, the two sets in (3.3) are

$$W_{(\gamma_1',\gamma_2')}(A)=(\gamma_1'-\gamma_2')W(B)+\tfrac{1}{2}(\gamma_1'+\gamma_2')\{\operatorname{tr}A\} \tag{3.4a}$$

and

$$W_{(\gamma_1,\gamma_2)}(A)=(2\alpha-1)(\gamma_1'-\gamma_2')W(B)+\tfrac{1}{2}(\gamma_1'+\gamma_2')\{\operatorname{tr}A\}, \tag{3.4b}$$

where $B=A-\frac{1}{2}(\operatorname{tr}A)I$.

It is known (e.g., [1], Sec. 166) that the numerical range of any 2×2 matrix is an ellipse (possibly degenerate) with the eigenvalues as foci. That is, $W(B)$ is an ellipse centered at $\frac{1}{2}\operatorname{tr}B$. In our case $\operatorname{tr}B=0$, so $(\gamma_1'-\gamma_2')W(B)$ is convex and symmetric with respect to the origin. Therefore, since $-1\leqslant 2\alpha-1\leqslant 1$, we have

$$(2\alpha-1)(\gamma_1'-\gamma_2')W(B)\subset(\gamma_1'-\gamma_2')W(B).$$

Hence the set in (3.4a) includes the set in (3.4b), and (3.3) follows. ■

LEMMA 3. *If c is obtained from c' by pinching, then*

$$W_c(A)\subset W_{c'}(A) \qquad \forall A\in\mathbf{C}_{n\times n}. \tag{3.5}$$

Proof. Let i,j, $i<j$, be the pinching indices described in (2.1). Every fixed choice of $n-2$ orthonormal vectors in $\mathbf{C}^n$,

$$x_1,\dots,x_{i-1},x_{i+1},\dots,x_{j-1},x_{j+1},\dots,x_n, \tag{3.6}$$

determines a 2-space, $\mathbf{X}$, perpendicular to these vectors. The values of $W_c(A)$

and $W_{c'}(A)$ corresponding to the vectors in (3.6) are, respectively,

$$\sum_{\substack{k=1\\k\neq i,j}}^{n} \gamma_k(Ax_k,x_k)+W_{(\gamma_i,\gamma_j)}(PA) \tag{3.7a}$$

and

$$\sum_{\substack{k=1\\k\neq i,j}} \gamma_k(Ax_k,x_k)+W_{(\gamma_i',\gamma_j')}(PA). \tag{3.7b}$$

Here P is the projection of $\mathbf{C}^n$ on $\mathbf{X}$, and it is understood that $W_{(\alpha,\beta)}(PA)$ is defined over $\mathbf{X}$, i.e.,

$$W_{(\alpha,\beta)}(PA)=\{\alpha(Ax,x)+\beta(Ay,y):x,y\in\mathbf{X};\ x,y \text{ orthonormal}\}.$$

Since $\mathbf{X}$ is 2-dimensional and PA maps $\mathbf{X}$ into itself, the restriction of PA to $\mathbf{X}$ may be presented by a 2×2 matrix. Moreover, it is clear from (2.1) that since c is a pinch of c', (γ_i,γ_j) is obtained from (γ_i',γ_j') by the same pinching. Thus, Corollary 1 implies that

$$W_{(\gamma_i,\gamma_j)}(PA)\subset W_{(\gamma_i',\gamma_j')}(PA).$$

Consequently, the set in (3.7b) includes the set of (3.7a). Since the vectors in (3.6) were arbitrary, the relation (3.5) holds, and the proof is complete. ∎

The following theorem is an immediate consequence of Lemma 3.

THEOREM 3. *If $c \prec\prec c'$, then*

$$W_c(A)\subset W_{c'}(A)\qquad \forall A\in\mathbf{C}_{n\times n}. \tag{3.8}$$

Proof. By hypothesis, there exists a finite sequence, $c'=c_1,c_2,\dots,c_l=c$, such that each c_i $(1<i\leqslant l)$ is obtained from c_{i-1} by pinching. So, by Lemma 3,

$$W_c(A)=W_{c_l}(A)\subset\cdots\subset W_{c_1}(A)=W_{c'}(A)\qquad \forall A\in\mathbf{C}_{n\times n},$$

and (3.8) follows. ∎

At this point it would be natural to ask whether $c\prec c'$ implies (3.8) or not. To answer this question in the negative take $A=\operatorname{diag}(0,1,i)$ and

$c'=(0,1,i)$. Westwick [2], has shown that $W_{c'}(A)$ includes the points 1 and $2i$, but not the open line segment joining them. In particular, $(1+2i)/2\notin W_{c'}$. Now take $c=(\frac{1}{2},i/2,\frac{1}{2}+i/2)$. By (2.2) we have that $c\prec c'$; yet the point

$$\gamma_1(Ae_3,e_3)+\gamma_2(Ae_1,e_1)+\gamma_3(Ae_2,e_2)=\frac{1+2i}{2}$$

of $W_c(A)$ does not belong to $W_{c'}(A)$.

A somewhat weaker result holds for the relation $\prec$, and we establish first the next lemma.

LEMMA 4. *Given two bounded disjoint convex sets $\mathcal{K}_1$, $\mathcal{K}_2$ in $\mathbf{C}^n$, there exists a linear functional φ on $\mathbf{C}^n$ such that $\varphi(x)\neq\varphi(y)$ for all $x\in\mathcal{K}_1$, $y\in\mathcal{K}_2$.*

Proof. We first consider $\mathcal{K}_1$, $\mathcal{K}_2$ as convex sets in $\mathbf{R}^{2n}$. By the separation theorem for real vector spaces (e.g., [5], Theorem 20, p. 204), there exists a linear real functional $\psi(x)$ on $\mathbf{R}^{2n}$ such that $\psi(x)<\psi(y)$ for all $x\in\mathcal{K}_1$, $y\in\mathcal{K}_2$. More explicitly we have

$$\psi(x)=\beta_{11}\xi_{11}+\beta_{12}\xi_{12}+\beta_{21}\xi_{21}+\beta_{22}\xi_{22}+\cdots+\beta_{n1}\xi_{n1}+\beta_{n2}\xi_{n2},$$

where $x=(\xi_1,\dots,\xi_n)$, $\xi_j=\xi_{j1}+i\xi_{j2}$, and the β_{ij} are real coefficients. Now define a complex functional on $\mathbf{C}^n$:

$$\varphi(x)=\beta_1\xi_1+\cdots+\beta_n\xi_n,\qquad \beta_j=\beta_{j1}-i\beta_{j2}.$$

It is easily seen that $\psi(x)=\operatorname{Re}\varphi(x)$; so $\operatorname{Re}\varphi(x)<\operatorname{Re}\varphi(y)$ for $x\in\mathcal{K}_1$, $y\in\mathcal{K}_2$, and the lemma follows. ∎

THEOREM 4. *We have $c\prec c'$ if and only if*

$$W_c(A)\subset\operatorname{conv}\{W_{c'}(A)\}\qquad\forall A\in\mathbf{C}_{n\times n}. \tag{3.9}$$

Proof. If $c\prec c'$, then for some doubly stochastic S we have $c=Sc'$. The matrix S is a convex combination of permutation matrices P_σ. Thus $c=\sum_\sigma\alpha_\sigma P_\sigma c'$, and the relation among the components of c and c' is

$$\gamma_j=\sum_\sigma\alpha_\sigma\gamma'_{\sigma(j)},\qquad j=1,\dots,n.$$

This yields that any point $\sum \gamma_j(Ax_j,x_j)$ of $W_c(A)$ satisfies

$$\sum \gamma_j(Ax_j,x_j) = \sum_{j=1}^{n} (Ax_j,x_j) \sum_{\sigma} \alpha_\sigma \gamma'_{\sigma(j)}$$

$$= \sum_{\sigma} \alpha_\sigma \left[\sum_{j=1}^{n} \gamma_{\sigma(j)}(Ax_j,x_j) \right].$$

That is, each point in W_c is a convex combination of points in $W_{c'}$, and (3.9) follows.

For the necessity part of the proof we recall that the condition $c \prec c'$ is equivalent to the fact that c belongs to the convex set

$$\mathcal{K}_1 = \{ Sc' : S \text{ doubly stochastic} \}.$$

Let $\mathcal{K}_2$ be the set which consists only of c. If $c \not\prec c'$, then $\mathcal{K}_1 \cap \mathcal{K}_2 = \varnothing$, and by Lemma 4 there exist complex coefficients $\beta_1,\ldots,\beta_n$ such that the linear functional $\varphi(x) = \sum_i \beta_i \xi_i$ satisfies

$$\varphi(c) \notin \{ \varphi(x) : x \in \mathcal{K}_1 \} = \{ \varphi(Sc') : S \text{ doubly stochastic} \}. \tag{3.10}$$

Consider now the matrix $B = \operatorname{diag}(\beta_1,\ldots,\beta_n)$. We have

$$\varphi(c) = \sum_{j=1}^{n} \beta_j \gamma_j = \sum \gamma_j (Be_j,e_j) \in W_c(B). \tag{3.11}$$

On the other hand, take any point $\sum_j \gamma'_j(Bx_j,x_j)$ in $W_{c'}(B)$. Here $\{x_j = (\xi_{1j},\ldots,\xi_{nj})\}_{j=1}^{n}$ is an orthonormal system in $\mathbf{C}^n$, so $\sum_i |\xi_{ij}|^2 = \sum_j |\xi_{ij}|^2 = 1$, and consequently the matrix X with $X_{ij} = |\xi_{ij}|^2$ is doubly stochastic. Hence

$$\sum_{j=1}^{n} \gamma'_j (Bx_j,x_j) = \sum_{j=1}^{n} \gamma'_j \sum_{i=1}^{n} \beta_i |\xi_{ij}|^2$$

$$= \sum_{i=1}^{n} \beta_i \sum_{j=1}^{n} X_{ij}\gamma'_j = \varphi(Xc').$$

This gives

$$W_{c'}(B) \subset \{ \varphi(Sc') : S \text{ doubly stochastic} \},$$

and since the set on the right side is convex, we get in fact

$$\operatorname{conv} W_{c'}(B) \subset \{\varphi(Sc') : S \text{ doubly stochastic}\}. \tag{3.12}$$

The inclusion in (3.12) together with (3.10), (3.11) yields $W_c(B) \not\subset \operatorname{conv} W_{c'}(B)$, and (3.9) is violated. ∎

4. THE CASE OF REAL COEFFICIENTS

For real vectors c the situation is much simpler. As in the complex case, the set $W_c(A)$ remains unchanged under permutations of the γ_j. Therefore, given a set of coefficients $\{\gamma_1, \dots, \gamma_n\}$, it will often be convenient to arrange them in decreasing order.

DEFINITION 2. A real vector $c = (\gamma_1, \dots, \gamma_n)$ is called *ordered* if

$$\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_n.$$

The convenience of ordering real vectors is demonstrated in the next lemma.

LEMMA 5. *If c' is ordered and $c \prec c'$, then*

$$\sum_{j=1}^{k} \gamma_j \leqslant \sum_{j=1}^{k} \gamma_j', \qquad k = 1, \dots, n, \tag{4.1}$$

with equality for $k = n$.

Proof. If $c \prec c'$, then for some doubly stochastic S, we have $c = Sc'$. Hence for a fixed k, $1 \leqslant k \leqslant n$,

$$\sum_{i=1}^{k} \gamma_i = \sum_{i=1}^{k} \sum_{j=1}^{n} S_{ij} \gamma_j' = \sum_{j=1}^{n} \left(\sum_{i=1}^{k} S_{ij} \right) \gamma_j'. \tag{4.2}$$

Setting

$$\alpha_j = \sum_{i=1}^{k} S_{ij}, \qquad j = 1, 2, \dots, n,$$

we have

$$0 \leqslant \alpha_j \leqslant 1 \quad \text{and} \quad \sum_{j=1}^{n} \alpha_j = k. \tag{4.3}$$

So, using the fact that c' is ordered, we get from (4.2), (4.3)

$$\sum_{i=1}^{k} \gamma_i = \sum_{j=1}^{n} \alpha_j \gamma_j' \leqslant \sum_{j=1}^{k} \gamma_j'.$$

For $k = n$ each $\alpha_j = 1$ and we have equality. ∎

We remark that the relations in (4.1) are discussed in Chapter 2 of [6], beginning with Sec. 2.18.

Two more preliminary results lead to Theorem 5.

LEMMA 6. *Let γ_i', γ_j' with $\gamma_i' > \gamma_j'$ be two real components of c'. Let δ satisfy $0 \leqslant \delta \leqslant \gamma_i' - \gamma_j'$. Then*

$$c \equiv c' - \delta(e_i - e_j)$$

is a pinch of c'.

Proof. Denote $\alpha' = \delta/(\gamma_i' - \gamma_j')$. Evidently $0 \leqslant \alpha' \leqslant 1$, and by the definition of c we have

$$\gamma_i = \gamma_i' - \delta = \gamma_i' - \alpha'(\gamma_i' - \gamma_j') = (1-\alpha')\gamma_i' + \alpha'\gamma_j' \tag{4.4a}$$

and

$$\gamma_j = \gamma_j' + \delta = \gamma_j' + \alpha'(\gamma_i' - \gamma_j') = \alpha'\gamma_i' + (1-\alpha')\gamma_j'. \tag{4.4b}$$

Equations (4.4) are equivalent to (2.1); hence c is a pinch of c' and the statement is proven. ∎

LEMMA 7. *Let c, c' be ordered. If c, c' satisfy* (4.1) *with equality for $k = n$, then $c \prec\prec c'$*

Proof. The idea of the proof is to construct a sequence of vectors

$c'=c_1,c_2,\ldots,$ such that each c_i has the following three properties. First,

$$c_{i+1} \prec\prec c_i, \qquad i \geqslant 1; \tag{4.5}$$

second,

$$\sum_{j=1}^{k} \gamma_j \leqslant \sum_{j=1}^{k} \gamma_{ij}, \qquad k=1,2,\ldots,n, \tag{4.6}$$

with equality for $k=n$; and third, the number of equal elements in the sets $\{\gamma_{i1},\ldots,\gamma_{in}\}$ and $\{\gamma_1,\ldots,\gamma_n\}$ is at least $i-1$. Here the γ_j and γ_{ij} are, respectively, the components of c and c_i.

By the last two properties, there exists a finite l ($l \leqslant n$), for which $c_l = c$. Hence, by property (4.5) we get

$$c=c_l \prec\prec \cdots \prec\prec c_1 = c', \tag{4.7}$$

which leads by transitivity to the desired result $c \prec\prec c'$.

As indicated, we start by choosing $c_1 = c'$, for which the first and third properties are satisfied in a trivial manner, and the second, by the hypothesis.

Now suppose that $c_1,\ldots,c_i$ with the above properties has been constructed. If $c_i = c$, then the sequence (4.7) is complete; so let us assume $c_i \neq c$ and construct c_{i+1}. We have the inequalities in (4.6), from which we conclude that there exists an r, $1 \leqslant r < n$, such that

$$\gamma_1 = \gamma_{i1},\ldots,\gamma_{r-1} = \gamma_{i,r-1}; \qquad \gamma_r < \gamma_{ir}, \tag{4.8a}$$

and a least s, $r < s \leqslant n$, such that

$$\gamma_s > \gamma_{is}. \tag{4.8b}$$

Since c is ordered, we have $\gamma_r \geqslant \gamma_s$, which together with (4.8) gives $\gamma_{ir} > \gamma_r \geqslant \gamma_s > \gamma_{is}$. So the quantity

$$\delta = \min\{\gamma_{ir} - \gamma_r, \gamma_s - \gamma_{is}\} \tag{4.9}$$

satisfies $0 < \delta < \gamma_{ir} - \gamma_{is}$. Hence, by Lemma 6,

$$c_{i+1} \equiv c_i - \delta(e_r - e_s) \tag{4.10}$$

is a pinch of c_i. So $c_{i+1} \prec\prec c_i$, i.e., c_{i+1} has the first property (4.5).

Next, we wish to show that c_{i+1} has the second property, that is

$$\sum_{j=1}^{k} \gamma_i \leqslant \sum_{j=1}^{k} \gamma_{i+1,j}, \qquad k=1,\ldots,n,$$

with equality for $k=n$. Since c_i satisfies (4.6), and since c_{i+1} is obtained from c_i by changing only the r and s components while their sum is preserved, it is clear that for any k with $1 \leqslant k < r$ or $s \leqslant k \leqslant n$, we have

$$\sum_{j=1}^{k} \gamma_j \leqslant \sum_{j=1}^{k} \gamma_{ij} = \sum_{j=1}^{k} \gamma_{i+1,j}.$$

Now use (4.9)–(4.10) to find that

$$\gamma_{i+1,r} = \gamma_{ir} - \delta \geqslant \gamma_{ir} - (\gamma_{ir} - \gamma_r) = \gamma_r.$$

So

$$\sum_{j=1}^{k} \gamma_j \leqslant \sum_{j=1}^{k} \gamma_{ij} \leqslant \sum_{j=1}^{k} \gamma_{i+1,j}$$

for $r \leqslant k < s$ also.

Finally, consider the third property. According to the construction of c_{i+1}, we have $\gamma_{i+1,r} = \gamma_r$ or $\gamma_{i+1,s} = \gamma_s$, or both. So, by comparing with (4.8), we see that the number of components of c_{i+1} which equal components of c is greater than the number of equalities for c_i and c, and is therefore at least i. This completes the proof. ∎

Combining Lemmas 5 and 7, together with Theorem 1, we easily obtain the following.

THEOREM 5. *Let c, c' be ordered vectors. Then each of the relations $c \prec c'$ and $c \prec\prec c'$ is equivalent to*

$$\sum_{j=1}^{k} \gamma_j \leqslant \sum_{j=1}^{k} \gamma'_j, \qquad k=1,\ldots,n, \tag{4.11}$$

with equality for $k=n$.

In general, it is more convenient to verify the condition (4.11) than to check whether $c \prec c'$ or $c \prec\prec c'$ according to the original definitions.

Since the relations $c \prec c'$ and $c \prec\prec c'$ are preserved under permutations of the γ_j, γ'_j, we rephrase part of Theorem 5:

THEOREM 6. *If c, c' are real vectors, then the relations $c \prec c'$ and $c \prec\prec c'$ are equivalent.*

We come now to one of the main results.

THEOREM 7. *If c, c' are real, then $c \prec c'$ if and only if*

$$W_c(A) \subset W_{c'}(A) \qquad \forall A \in \mathbf{C}_{n\times n}. \tag{4.12}$$

Proof. By Theorem 6, $c \prec c'$ implies $c \prec\prec c'$, so by Theorem 3 we have (4.12). Conversely, (4.12) yields (3.9) and by Theorem 4, $c \prec c'$. ∎

REMARK. Theorem 7 can be obtained immediately from Theorem 4, using the fact that for real c, W_c is convex, i.e., $W_c = \operatorname{conv}\{W_c\}$. Yet, the convexity if W_c is not essential to the proof.

COROLLARY 2.

(a) *If $c = (\gamma_1, \dots, \gamma_n) \in \mathbf{R}^n$ with $\sum_i \gamma_i = \alpha$, then $(\alpha/n, \dots, \alpha/n) \prec c$, and hence*

$$\left\{\frac{\alpha}{n}\operatorname{tr} A\right\} \subset W_c(A) \qquad \forall A \in \mathbf{C}_{n\times n}. \tag{4.13}$$

(b) *If $\gamma_i \geqslant 0$, then $c \prec (\alpha, 0, \dots, 0)$ and*

$$W_c(A) \subset \alpha W(A) \qquad \forall A \in \mathbf{C}_{n\times n}.$$

(c) *If $\alpha = 0$ then*

$$\bigcap_{A \in \mathbf{C}_{n\times n}} W_c(A) = \{0\}. \tag{4.14}$$

Proof. First take the ordered version of c and observe that Theorems 5 and 7 yield (a) and (b). Now, if $\alpha = 0$, then according to (4.13), $0 \in W_c(A)$ for all A. Since $W_c(0) = \{0\}$, we have (4.14) and the corollary follows.

Note that (b) follows directly from the convexity of $W(A)$. ∎

COROLLARY 3 (Fillmore and Williams). *The k-numerical ranges satisfy*

$$\left\{\frac{1}{n}\operatorname{tr} A\right\} = W_n(A) \subset \cdots \subset W_2(A) \subset W_1(A) \equiv W(A). \tag{4.15}$$

Proof. By (1.1), $W_s(A) = W_{c_s}(A)$ with the ordered vector

$$c_s = (\gamma_{s1}, \dots, \gamma_{sn}) = \frac{1}{s}(e_1 + \cdots + e_s).$$

For each s, $1 \leqslant s < n$ we have

$$\sum_{j=1}^{k} \gamma_{sj} = \frac{1}{s}\min\{k,s\} \geqslant \frac{1}{s+1}\min\{k,s+1\} = \sum_{j=1}^{k} \gamma_{s+1,j},$$

with equality for $k = n$. So Theorem 5 implies that $c_{s+1} \prec c_s$, $1 \leqslant s < n$. Hence, by Theorem 7,

$$W_{s+1}(A) = W_{c_{s+1}}(A) \subset W_{c_s}(A) = W_s(A), \qquad s = 1, \dots, n-1,$$

and (4.15) holds. ■

This result was obtained in a different way, using the convexity of W_k, by Fillmore and Williams [7].

Remark. In general, for given vectors $c = (\gamma_1, \dots, \gamma_n)$, $c' = (\gamma_1', \dots, \gamma_n')$, there exists no constant λ such that $c \prec \lambda c'$. To demonstrate this statement assume that c, c' are ordered and that $\sum \gamma_j' > 0$, $\sum \gamma_j \geqslant 0$. If $c \prec \lambda c'$, then for some doubly stochastic S we would have $c = \lambda S c'$, which yields

$$\sum_{i=1}^{n} \gamma_i = \lambda \sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij}\gamma_j' = \lambda \sum_{j=1}^{n} \gamma_j'.$$

Consequently

$$\lambda = \sum \gamma_i / \sum \gamma_j', \tag{4.16}$$

so $\lambda \geqslant 0$, and $\lambda c'$ is ordered. Now, by Theorem 5 we should have

$$\sum_{j=1}^{k} \gamma_j \leqslant \lambda \sum_{j=1}^{k} \gamma_j', \qquad k = 1, \dots, n, \tag{4.17}$$

with equality for $k = n$. But as λ of (4.16) satisfies (4.17) for $k = n$, it will not, in general, satisfy the rest of (4.17).

The situation is quite different in the homogeneous case $\Sigma\gamma_j = \Sigma\gamma_j' = 0$, where we have the following result.

LEMMA 8. *Let c, c' be ordered vectors with $\Sigma_j\gamma_j = \Sigma_j\gamma_j' = 0$ and $c' \neq 0$. Set*

$$\eta = \eta(c,c') = \max_{1 \leqslant k < n} \frac{\gamma_1 + \cdots + \gamma_k}{\gamma_1' + \cdots + \gamma_k'}, \tag{4.18a}$$

$$\zeta = \zeta(c,c') = \min_{1 \leqslant k < n} \frac{\gamma_1 + \cdots + \gamma_k}{\gamma_n' + \cdots + \gamma_{n-k+1}'}. \tag{4.18b}$$

Then $c \prec \lambda c'$ if and only if $\lambda \geqslant \eta$ or $\lambda \leqslant \zeta$.

Proof. First we show that

$$\gamma_1' + \cdots + \gamma_k' > 0, \quad \gamma_n' + \cdots + \gamma_{n-k+1}' < 0; \qquad k = 1, \dots, n-1.$$

Since $\Sigma\gamma_j' = 0$, it suffices to prove the left inequalities, so assume that $\gamma_1' + \cdots + \gamma_k' \leqslant 0$ for some $k < n$. This means that $\gamma_{k+1}' + \cdots + \gamma_n' \geqslant 0$, thus $\gamma_{k+1}' \geqslant 0$, and consequently $\gamma_1' \geqslant \cdots \geqslant \gamma_k' \geqslant \gamma_{k+1}' \geqslant 0$. Since $c' \neq 0$, we have $\gamma_1' > 0$ and our assumption is contradicted. Similarly, the partial sums $\gamma_1 + \cdots + \gamma_k$, $k < n$, are non-negative, and it follows that η, ζ of (4.18) are well defined and satisfy $\eta \geqslant 0$, $\zeta \leqslant 0$.

Now choose λ with $\lambda \geqslant 0$. The vector $\lambda c'$ remains ordered, and according to Theorem 5, $c \prec \lambda c'$ if and only if

$$\lambda \sum_{j=1}^{k} \gamma_j' \geqslant \sum_{j=1}^{k} \gamma_j, \qquad k = 1, \dots, n,$$

with equality for $k = n$. The hypothesis $\Sigma\gamma_j = \Sigma\gamma_j' = 0$ implies equality for $k = n$; so $c \prec \lambda c'$ is equivalent to

$$\lambda \sum_{j=1}^{k} \gamma_j' \geqslant \sum_{j=1}^{k} \gamma_j, \qquad k = 1, \dots, n-1. \tag{4.19}$$

However, by the definition of η,

$$\eta \sum_{j=1}^{k} \gamma_j' \geqslant \sum_{j=1}^{k} \gamma_j, \qquad k = 1, \dots, n-1,$$

with equality for some $1 \leqslant k < n$. Thus, (4.19) holds if and only if $\lambda \geqslant \eta$.

If $\lambda<0$, then $\lambda c'$ becomes unordered, and its equivalent ordered version with a positive multiplier is $(-\lambda)(-\gamma_n',\ldots,-\gamma_1')$. Using the previous argument, we find that $c\prec\lambda c'$ if and only if

$$-\lambda\geqslant\max_{1\leqslant k<n}\frac{\gamma_1+\cdots+\gamma_k}{-\gamma_n'-\cdots-\gamma_{n-k+1}'}=-\min_{1\leqslant k<n}\frac{\gamma_1+\cdots+\gamma_k}{\gamma_n'+\cdots+\gamma_{n-k+1}'}=-\zeta,$$

and the lemma follows. ■

Theorem 7 and Lemma 8 have an immediate consequence.

THEOREM 8. *Let c, c' be ordered vectors with $\Sigma_j\gamma_j=\Sigma_j\gamma_j'=0$ and $c'\neq 0$. Then*

$$W_c(A)\subset W_{\lambda c'}(A)\equiv\lambda W_{c'}(A)\qquad\forall A\in\mathbf{C}_{n\times n}$$

if and only if $\lambda\geqslant\eta(c,c')$ or $\lambda\leqslant\zeta(c,c')$, where η, ζ are defined in (4.18).

This result is obtained differently in [8].

COROLLARY 4. *Let $a=(\alpha_1,\ldots,\alpha_n)$ and $a'=(\alpha_1',\ldots,\alpha_n')$ be ordered vectors such that not all the components of a' are equal. Set $\alpha=\Sigma\alpha_j$, $\alpha'=\Sigma\alpha_j'$, and define*

$$c=a-(\alpha/n,\ldots,\alpha/n),\qquad c'=a'-(\alpha'/n,\ldots,\alpha'/n).$$

Then

$$W_a(A)-\left\{\frac{\alpha}{n}\operatorname{tr}A\right\}\subset\lambda\left(W_{a'}(A)-\left\{\frac{\alpha'}{n}\operatorname{tr}A\right\}\right)\qquad\forall A\in\mathbf{C}_{n\times n}$$

if and only if $\lambda\geqslant\eta(c,c')$ or $\lambda\leqslant\zeta(c,c')$, where η, ζ are given in (4.18).

Proof. The components of the vectors c, c' satisfy $\Sigma\gamma_j=\Sigma\gamma_j'=0$, and $c'\neq 0$. Hence, by Theorem 8,

$$W_a(A)-\left\{\frac{\alpha}{n}\operatorname{tr}A\right\}=W_c(A)\subset\lambda W_{c'}(A)$$

$$=\lambda\left(W_{a'}(A)-\frac{\alpha'}{n}\operatorname{tr}A\right)\qquad\forall A\in\mathbf{C}_{n\times n}$$

if and only if the conditions of the corollary are satisfied. ■

5. GENERALIZED NUMERICAL RADIUS

A concept which directly relates to the generalized numerical range $W_c(A)$ is the generalized numerical radius

$$r_c(A) = \max\{|z| : z \in W_c(A)\}$$

$$= \max\left\{\left|\sum_{j=1}^{n} \gamma_j (Ax_j, x_j)\right| : (x_1, \ldots, x_n) \in \Lambda_n\right\}.$$

In particular we have the k-numerical radius

$$r_k(A) = \max\{|z| : z \in W_k(A)\}, \qquad k = 1, 2, \ldots, n,$$

which reduces, for $k = 1$, to the classical numerical radius

$$r(A) = \max\{|z| : z \in W(A)\} = \max_{|x|=1} |(Ax, x)|.$$

The function $r(A)$ provides an important tool in the linear stability analysis of multidimensional hyperbolic and parabolic initial value problems (e.g., [9], Sec. 2), and one may expect that the generalized radius will be applicable as well.

It is obvious that if $W_c(A) \subset W_{c'}(A)$ or even if $W_c(A) \subset \operatorname{conv} W_{c'}(A)$, then $r_c(A) \leqslant r_{c'}(A)$, though the converse may fail to hold. Thus, we use Theorems 4, 8 and Corollaries 2, 3 to obtain, respectively, the following results.

THEOREM 9.

(a) *If c, c' are complex n-vectors with $c \prec c'$, then*

$$r_c(A) \leqslant r_{c'}(A) \qquad \forall A \in \mathbf{C}_{n \times n}. \tag{5.1}$$

(b) *Let c, c' be real ordered vectors with $\Sigma \gamma_j = \Sigma \gamma'_j = 0$ and $c' \neq 0$. Let λ satisfy $\lambda \geqslant \eta(c, c')$ or $\lambda \leqslant \zeta(c, c')$, where η, ζ are defined in* (4.18). *Then*

$$r_c(A) \leqslant |\lambda| r_{c'}(A) \qquad \forall A \in \mathbf{C}_{n \times n}.$$

(c) *For $c = (\gamma_1, \ldots, \gamma_n)$ real with $\Sigma_j \gamma_j = \alpha$,*

$$\frac{|\alpha|}{n} |\operatorname{tr} A| \leqslant r_c(A) \qquad \forall A \in \mathbf{C}_{n \times n}.$$

If $\gamma_j \geqslant 0$, *then*

$$r_c(A) \leqslant \alpha r(A) \qquad \forall A \in \mathbf{C}_{n\times n}.$$

(d) *The k-numerical radii satisfy*

$$\frac{1}{n}|\operatorname{tr} A| = r_n(A) \leqslant \cdots \leqslant r_1(A) = r(A) \qquad \forall A \in \mathbf{C}_{n\times n}.$$

6. C-NUMERICAL RANGES

The numerical ranges defined in this paper can be generalized in the following way.

DEFINITION 3. Let $C \in \mathbf{C}_{n\times n}$ be fixed, and let $\mathfrak{U}_n$ denote the group of $n \times n$ unitary matrices. We call the set

$$W_C(A) = \{\operatorname{tr}(CU^*AU) : U \in \mathfrak{U}_n\}$$

the *C-numerical range* of the n-square matrix A.

If $c = (\gamma_1, \dots, \gamma_n)$ is a given vector, we take $D = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$ and find that

$$\begin{aligned} W_c(A) &= \left\{ \sum_{j=1}^{n} \gamma_j x_j^* A x_j : (x_1, \dots, x_n) \in \Lambda_n \right\} \\ &= \{\operatorname{tr}(DU^*AU) : U \in \mathfrak{U}_n\} = W_D(A). \end{aligned} \tag{6.1}$$

So, indeed, $W_c(A)$ is a special case of the C-numerical range. In fact our last result will characterize the class of matrices C for which there exists a vector c such that

$$W_C(A) = W_c(A) \qquad \forall A \in \mathbf{C}_{n\times n}.$$

First, we give two simple properties of the C-numerical range.

LEMMA 9.

(a) *For any* $C, A \in \mathbf{C}_{n\times n}$ *we have*

$$W_C(A) = W_A(C). \tag{6.2}$$

(b) *The set $W_C(A)$ is invariant under unitary similarities of C or of A.*

Proof. We have

$$W_C(A) = \{\mathrm{tr}(CU^*AU) : U \in \mathcal{U}_n\} = \{\mathrm{tr}(U^*AUC) : U \in \mathcal{U}_n\}$$

$$= \{\mathrm{tr}(AUCU^*) : U \in \mathcal{U}_n\} = W_A(C),$$

so (6.2) holds, and it follows that C and A play a symmetric role in the definition of $W_C(A)$. Hence, for part (b), it suffices to show that $W_C(A)$ is invariant under unitary similarities of A. But that is an immediate consequence of Definition 3 which states that $W_C(A)$ depends only on the class, $\mathcal{S}(A) = \{U^*AU : U \in \mathcal{U}_n\}$, of matrices unitarily similar to A. ∎

The next result leads to Theorem 10.

LEMMA 10. *If $\mathcal{S}$, $\mathcal{S}'$ are compact connected subsets of $\mathbf{C}_{n\times n}$ such that*

$$\{\varphi(X) : X \in \mathcal{S}\} = \{\varphi(X') : X' \in \mathcal{S}'\} \tag{6.3}$$

for all linear functionals φ on $\mathbf{C}_{n\times n}$, then

$$\mathcal{K} \equiv \mathrm{conv}\{\mathcal{S}\} = \mathrm{conv}\{\mathcal{S}'\} \equiv \mathcal{K}'.$$

Proof. We recall that the hyperplanes (of real dimension $2n^2-1$) of $\mathbf{C}_{n\times n}$ are the loci of the equations

$$\mathrm{Re}\,\varphi(X) = \alpha$$

as φ varies over the nonzero functionals in $\mathbf{C}^*_{n\times n}$ and α varies in $\mathbf{R}$. Since $\mathcal{S}$ is connected, a hyperplane intersects $\mathcal{S}$ if and only if it intersects conv$\{\mathcal{S}\}$; thus (6.3) implies

$$\{\mathrm{Re}\,\varphi(X) : X \in \mathcal{K}\} = \{\mathrm{Re}\,\varphi(X') : X' \in \mathcal{K}'\} \qquad \forall \varphi \in \mathbf{C}^*_{n\times n}. \tag{6.4}$$

Now choose a functional φ and consider the set of real values

$$\mathcal{R}_\varphi(\mathcal{K}) = \{\mathrm{Re}\,\varphi(X) : X \in \mathcal{K}\}.$$

Since $\mathcal{K}$ is compact and connected, $\mathcal{R}_\varphi(\mathcal{K})$ is a closed interval with end points

$$\mu_1 = \min \mathcal{R}_\varphi(\mathcal{K}), \qquad \mu_2 = \max \mathcal{R}_\varphi(\mathcal{K}).$$

This means that a hyperplane $\operatorname{Re}\varphi(X)=\alpha$ intersects $\mathcal{K}$ if and only if $\alpha\in\mathcal{R}_\varphi(\mathcal{K})$, and in particular

$$\operatorname{Re}\varphi(X)=\mu_1,\qquad \operatorname{Re}\varphi(X)=\mu_2 \tag{6.5}$$

are the two planes of support for $\mathcal{K}$, defined by φ.

According to (6.4)

$$\mathcal{R}_\varphi(\mathcal{K})=\mathcal{R}_\varphi(\mathcal{K}')\qquad \forall\varphi\in\mathbf{C}^*_{n\times n}; \tag{6.6}$$

so the hyperplanes in (6.5) support $\mathcal{K}'$ as well as $\mathcal{K}$, for all φ. Since compact convex sets are uniquely determined by their supporting planes, the proof is complete. ∎

THEOREM 10. *We have*

$$W_C(A)=W_{C'}(A)\qquad \forall A\in\mathbf{C}_{n\times n} \tag{6.7}$$

if and only if C, C' *are unitarily similar.*

Proof. If C, C' are unitarily similar, then (6.7) is given by part (b) of Lemma 9.

For the converse we use (a) of Lemma 9, by which the hypothesis in (6.7) becomes $W_A(C)=W_A(C')$ for all A; or more explicitly

$$\{\operatorname{tr}(AU^*CU):U\in\mathcal{U}_n\}=\{\operatorname{tr}(AU^*C'U):U\in\mathcal{U}_n\}\qquad \forall A\in\mathbf{C}_{n\times n}. \tag{6.8}$$

Next we remember that every linear functional φ on $\mathbf{C}_{n\times n}$ is of the form

$$\varphi(X)=\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_{ij}\xi_{ji}=\operatorname{tr}(AX),$$

where $A=[\alpha_{ij}]$ is a matrix of coefficients, and $X=[\xi_{ij}]$ is arbitrary. Thus, the hypothesis in (6.8) takes the form

$$\{\varphi(X):X\in\mathcal{S}\}=\{\varphi(X'):X'\in\mathcal{S}'\}\qquad \forall\varphi\in\mathbf{C}^*_{n\times n},$$

where

$$\mathcal{S}=\{U^*CU:U\in\mathcal{U}_n\},\qquad \mathcal{S}'=\{U^*C'U:U\in\mathcal{U}_n\}$$

are compact connected subsets of $\mathbf{C}_{n\times n}$. Consequently, by Lemma 10,

$$\mathcal{K} \equiv \operatorname{conv}\{\mathcal{S}\} = \operatorname{conv}\{\mathcal{S}'\} \equiv \mathcal{K}'. \tag{6.9}$$

The sets $\mathcal{K}$, $\mathcal{K}'$ are compact, so they are spanned by the extreme points of $\mathcal{S}$ and $\mathcal{S}'$, respectively. Therefore, by the equality in (6.9) we finally get

$$\operatorname{ext}\{\mathcal{S}\} = \operatorname{ext}\{\mathcal{S}'\}.$$

Now take a point $U_1^*CU_1$ in $\operatorname{ext}\{\mathcal{S}\}$. It equals a point $U_2^*C'U_2$ in $\operatorname{ext}\{\mathcal{S}'\}$ where U_1, U_2 are both unitary. That is,

$$C = U^*C'U \qquad \text{with} \quad U = U_2U_1^*,$$

and the theorem is proven. ■

Our last result characterizes the relation between the C-numerical and the c-numerical ranges.

Corollary 5. *For a given $C \in \mathbf{C}_{n\times n}$, there exists a vector $c \in \mathbf{C}^n$ such that*

$$W_C(A) = W_c(A) \qquad \forall A \in \mathbf{C}_{n\times n} \tag{6.10}$$

if and only if C is normal. If C is normal, then the components of c are the eigenvalues of C in an arbitrary order.

Proof. By (6.1), the equality in (6.10) is equivalent to having a diagonal $D = \operatorname{diag}(\gamma_1, \dots, \gamma_n)$ such that

$$W_C(A) = W_D(A) \qquad \forall A \in \mathbf{C}_{n\times n}.$$

But C is unitarily similar to a diagonal matrix if and only if C is normal, so Theorem 10 completes the proof. ■

Note that if C is normal with real eigenvalues—that is, Hermitian—then (6.10) holds with a real c, and by Westwick's theorem $W_C(A)$ is convex.

We conclude this paper with the following discussion.

Remark. It is clear now that $W_C(A)$ is the range of values of the mapping

$$\varphi : \mathcal{S}(A) \to \mathbf{C},$$

where

$$\mathcal{S}(A)=\{U^*AU : U\in\mathcal{U}_n\}\subset\mathbf{C}_{n\times n},$$

and φ is the linear functional on $\mathbf{C}_{n\times n}$ defined by

$$\varphi(X)=\mathrm{tr}(CX).$$

That is, $W_C(A)$ gives us *all* the information a single functional can provide about the set $\mathcal{S}(A)$. From this point of view, $W_C(A)$ is an ultimate generalization of previous concepts of numerical ranges.

However, more information on $\mathcal{S}(A)$ could be obtained by considering mappings of the form

$$X\to(\varphi_1(X),\dots,\varphi_m(X))\in\mathbf{C}^m \qquad [X\in\mathcal{S}(A)],$$

where $\varphi_1,\dots,\varphi_m$ are functionals on $\mathbf{C}_{n\times n}$, and m is arbitrary. In fact we do not need $m>n^2$; for if we denote by φ_{ij} the functional defined by

$$\varphi_{ij}(X)=X_{ij}\equiv\xi_{ij},$$

then the mapping

$$X\to(\varphi_{11}(X),\dots,\varphi_{nn}(X))=(\xi_{11},\dots,\xi_{nn})\in\mathbf{C}^{n^2}$$

exactly characterizes the set $\mathcal{S}(A)$.

REFERENCES

1 P. R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, 1967.
2 R. Westwick, A theorem on numerical range, *Linear Multilinear Algebra* **2** (1975), 311–315.
3 M. Goldberg and E. G. Straus, Inclusion relations involving k-numerical ranges, *Linear Algebra Appl.*, **15** (1976), 261–270.
4 A. Horn and R. Steinberg, Eigenvalues of a unitary part of a matrix, *Pac. J. Math.* **9** (1959), 541–550.
5 H. L. Royden, *Real Analysis*, 2nd ed., Macmillan, 1968.
6 G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1952.
7 P. A. Fillmore and J. P. Williams, Some convexity theorems for matrices, *Glasg. Math. J.*, **12** (1971), 110–117.
8 M. Goldberg and E. G. Straus, Characterizations and integrals of generalized numerical ranges, *Pac. J. Math.*, to appear.
9 P. D. Lax and B. Wendroff, Difference schemes for hyperbolic equations with high order of accuracy, *Commun. Pure Appl. Math.* **XVII** (1964), 381–391.

Received 11 March 1976; revised 17 December 1976